**Analysis of Research Papers in Relation to Our Study**

I have analyzed the provided research papers and compared them to our study. Below is a detailed breakdown of **which research is most relevant, how our work is different, what we are lacking, and what gaps we are filling compared to existing research**.

**1. Research Closest to Our Study**

The **most relevant** research papers to our project are:

1. **Predictive Modeling of Hamstring Strain Injuries in Elite Australian Footballers**
   * This study used **machine learning models** (Random Forest, SVM, Neural Networks, Logistic Regression) to predict **hamstring injuries** based on eccentric hamstring strength, age, and previous injury history.
   * Used **SMOTE** for class imbalance handling.
   * Found that **ML models performed better than traditional statistical models** but had high variability in accuracy.
2. **A Machine Learning Approach to Assess Injury Risk in Elite Youth Football Players**
   * This study employed **XGBoost** to predict injury risks in **youth football players** based on **anthropometric, motor coordination, and physical performance measures**.
   * Used **SHAP analysis** to determine **feature importance**.
   * Found that **growth maturity and leg length** were strong injury predictors.
3. **A Preventive Model for Muscle Injuries Using Machine Learning**
   * Applied **decision trees, boosting techniques, and cost-sensitive learning** to predict **lower extremity muscle injuries**.
   * Found that **muscle imbalance, history of past injuries, poor sleep, and psychological burnout** were key predictors.
   * Showed that **muscle asymmetry is linked to a higher probability of injury**.

**2. How Our Study is More Advanced**

Compared to these research papers, **our study improves upon prior works in several ways**:

| **Feature** | **Existing Studies** | **Our Study** |
| --- | --- | --- |
| **Injury Type** | Focused mostly on **hamstring injuries** or broad muscle injuries. | Generalized approach for **various symmetry-related injuries** across **multiple symmetry metrics**. |
| **Feature Selection** | Used **manual feature selection** based on past research. | Used **feature importance ranking** with ML techniques (Random Forest importance, SHAP). |
| **Model Balancing Techniques** | Used **SMOTE** in some cases but did not explore **other resampling techniques**. | Compared **multiple resampling methods** (SMOTE, SMOTEENN, Class Weighting). |
| **Threshold-Based Categorization** | **Did not use thresholds dynamically**; mostly relied on **predefined cutoffs**. | Implemented a **dynamic buffer** approach based on **standard deviations** and **player-specific variations**. |
| **Athlete-Specific Risk Analysis** | Most models trained on **all players together**. | Categorized **athletes based on risk trends** across **multiple sessions**, enabling **individualized risk tracking**. |
| **Time-Series Analysis** | No research explicitly included **quarterly trends or time-based insights**. | We analyze **quarterly trends**, tracking how **risk categories** evolve over time. |
| **Real-World Validation** | Most models **were not validated on external datasets**. | Plan to **test our model on unseen 2025 athlete data** for generalization. |

**3. What Are We Lacking?**

Although our study is **more advanced in certain areas**, there are still some **drawbacks and potential improvements**:

| **Drawback** | **Improvement Needed** |
| --- | --- |
| **Lack of External Validation** | Need **external datasets from different teams/sports** to validate model performance. |
| **Limited Biomechanical Data** | Most features are from **force and torque symmetry**; adding **joint motion analysis, balance metrics, or GPS tracking** could improve accuracy. |
| **Psychological & Sleep Factors Ignored** | Studies show **poor sleep and burnout** contribute to injuries; including **sleep and psychological stress data** could enhance prediction. |
| **Real-Time Monitoring Not Considered** | If possible, integrating **wearable sensor data (heart rate, workload monitoring)** can improve real-time risk tracking. |

**4. What Gap Are We Filling?**

Our study **fills several key research gaps** that were **not** addressed in previous works:

1. **Using a Dynamic Buffer-Based Risk Threshold**
   * Instead of **predefined injury risk cutoffs**, we **calculate dynamic risk thresholds** based on **each athlete's deviation** from the population.
   * This allows for **more personalized risk detection**, avoiding **rigid one-size-fits-all rules** used in past research.
2. **Analyzing Injury Risk Over Time (Quarterly Trends)**
   * Unlike previous research, we **track how athlete risk levels evolve across multiple test sessions**.
   * This allows us to **detect worsening symmetry trends before injuries occur**.
3. **Combining Multiple Risk Handling Techniques**
   * We systematically compare **SMOTE, SMOTEENN, and Class Weighting** to **handle data imbalance**.
   * Previous research often relied **only on SMOTE** without testing **alternative resampling methods**.
4. **Feature Importance Analysis to Enhance Interpretability**
   * Most studies simply used **ML models without explaining why certain factors matter**.
   * We integrate **feature importance rankings (SHAP, Random Forest)** to **interpret why specific symmetry metrics matter most for injuries**.
5. **Bridging the Gap Between Injury Prediction & Actionable Insights**
   * Many past studies just **predict injury risk** without offering **clear prevention strategies**.
   * We aim to **provide actionable recommendations**, such as **training adjustments, strengthening routines, and intervention strategies** based on **high-risk athletes’ symmetry profiles**.

**5. Future Research Directions**

**How We Can Improve Our Study Further**

* **Expand Risk Factors Beyond Symmetry Metrics**
  + Add **joint kinematics (hip/knee movement tracking)**, **fatigue levels**, **reaction time**, and **psychological stress metrics**.
* **Validate Model Across Multiple Sports**
  + Test the model in **basketball, rugby, and athletics** to see if our findings hold across **different movement patterns**.
* **Real-Time Monitoring with Wearables**
  + Integrate **wearable sensors (IMU, GPS, smart insoles)** to detect **real-time imbalance trends** during matches.
* **Deep Learning for Video Analysis**
  + Use **computer vision models (PoseNet, OpenPose)** to **analyze athlete movement patterns and symmetry** in **match footage**.